Software Requirements Specification

for

TDH Data Requests Tool

Version 1.2

Nash Aragam

Office of Performance Management
Policy, Planning, and Assessment Division

September 15, 2014

Table of Contents

1. Overall Description

1.1 Purpose

The purpose of this document is to describe the functional specifications for a web-based application for Online Data Request Submission, Tracking, Metrics, and Response. This tool will be used to process requests for health-related data pertaining to the 95 counties of the state of Tennessee. This tool will reside within the Division of Policy, Planning, and Assessment (PPA) within the Tennessee Department of Health (TDH). PPA receives three common types of data requests: 1) requests for raw data only; 2) requests for the analysis, interpretation, and/or summary of data; and 3) requests for both raw data sets and the summary/analysis/interpretation. PPA has identified four tiers of data requesters:

- Tier 1 includes the Commissioner's Office, Governor's Office, Media, and Legislators/Congress persons;
- Tier 2 includes requests from the TDH's Office of General Counsel and Official Requests based on the Open Records and/or Freedom of Information Acts;
- Tier 3 includes internal (TDH) and external agency requests; and
- Tier 4 includes Academic and/or Qualified Researcher requests, as well as, requests from individuals of the general public.

PPA has a data request form which needs to be converted from a paper form layout to a web form layout. The objective is for PPA to have a web portal for any user to be able to register an account and submit a data request web form to launch PPA's workflow for registering the request, assigning the request, responding to the request and determining trends through business process analytics to conduct PPA process improvement. This tool is not for the purposes of transferring the data requested; this tool is for capturing the request itself and tracking its progression from data request registration through completion and any correspondence in between. Therefore, no protected health information (PHI) data will be transferred through this web application. The development of this tool will assist PPA with its business process improvement and increase the efficiency of staff resourcing to fulfill data requests. In addition, this tool will improve the requester satisfaction (quality) through improved interface and responsiveness with retrieving data from the State.

1.2 Project Scope

The Data Request is a part of the Data Reform activities that have been pursued by the Office of Performance Management (OPM) which resides in PPA of TDH. Currently, PPA has thirteen distinct datasets. Each dataset has an assigned data steward that serves as both the individual responsible for maintaining the dataset, as well as, serves as a supervisor to coordinate a response (via assigning his/her team member to respond to data requests submitted by Tier 1, 2, 3, and 4 categories of requesters). Data output provided in response to the data request is validated via an independent validation verification process.

The table below shows a brief inventory of the datasets that are used to fulfill the data requests submitted to PPA. The columns describe the data systems name, dataset name, years for which the data is available, the media on which the data resides, data location, data type (whether a flat file or a database), and some dataset details such as update frequency and the source of data. The blacked out column has some confidential access information controlled by the system administrator.

PPA DataSet Inventory - Primary Information										
Item #	Data Systems Name / Manager	Data Set Name	Years Available	Medium / Responsible Individual	Location & Type of the Data				Data Source Details	
					Database	FlatFile	Encrypted / Secure?		Update Frequency	Source(s)
1	Cancer Reporting Systems Data	Cancer Incidence Data (CID)	1983 - 2011	Electronic	MS Sequel Server	No	Yes		Daily	Hospitals, DMV[13] Records, Deaths (VS), ASTCs, Physicians, Pathology Labs
2	Vital Statistics Data	2.1 Births	1980 - 2012	Electronic	MS Access	Yes	No		Hourly	Mainframe (AIRS[14]) in Vital Records
		2.2 Deaths	1949 - 2012	Electronic	MS Access	Yes	No		Daily	KCC[15]
		2.3 Marriages	1978 - 2012		Mainframe AIRS?				Weekly	
		2.4 Divorces	1978 - 2012							
		2.5 Fetal Deaths	1960 - 2012	Electronic	MS Access	Yes	No		Bi-weekly	Manual entry by Hany
		2.6 Abortions (ITP)[3]	1975 - 2012						Weekly	KCC
		2.7 Birth/Death Matched	1993 - 2011		n/a	Yes	No		n/a	n/a
		2.8 Vital Statistics Bulletins	1927 - 2011	Electronic	n/a	No	No		Annually	n/a
3	Facilities Survey Data	3.1 Hospitals Individual JARs[4]	1986 - 1988, 1998 - 2009	Electronic	MS Access	No	No		Annually	Individual Hospitals around the State
			1966 - 2004, 2005 - 2013	(Paper / Other)						
		3.1.1 Hospitals Summary Reports (SRs)	1947 - 1974, 1980 - 1984, 1987 - 1997, 1999 - 2004, 2005 - 2013							
		3.2 Nursing Homes Individual JARs	1999 - 2001, 2003 - 2012	Electronic	MS Access, before 2009; Excel, 2009 and after	No	No		Annually	Individual Nursing Homes around the State
			1966 - 2003, 2003 - 2012	(Paper / Other)						
		3.2.1 Nursing Homes Summary Reports	1979 - 2012							
		3.3 Home Health Individual JARs	1999 - 2001, 2003 - 2013	Electronic	MS Access	No	No		Annually	Individual Home Healths around the State
			1978 - 2002, 2006 - 2013	(Paper / Other)						
		3.3.1 Home Health Summary Reports	1990 - 1996, 1998 - 2003, 2004 - 2005, 2006 - 2013							
		3.4 Hospices Individual JARs	2000 - 2013	Electronic	MS Access, before 2009; Excel, 2009 and after	No	No		Annually	Individual Hospices around the State
			1997 - 2004, 2005, 2006 - 2013	(Paper / Other)						
		3.4.1 Hospices Summary Reports	2004, 2005 - 2006, 2006 - 2013							
		3.5 ASTCs[5] Individual JARs	2000 - 2012	Electronic	MS Access, before 2009; Excel, 2009 and after	No	No		Annually	Individual ASTCs around the State
			2005, 2006 - 2012	(Paper / Other)						
		3.5.1 ASTCs Summary Reports	2004 - 2005, 2006 - 2009							
		3.6 ACLF[6] Individual JARs	2000 - 2001, 2003 - 2012	Electronic	MS Access, before 2009; Excel, 2009 and after	No	No		Annually	Individual ACLFs around the State
			1999 - 2004, 2005, 2006 - 2012	(Paper / Other)						
		3.6.1 ACLF Summary Reports	2004 - 2005, 2006 - 2009							
		3.7 ODC[7] Individual JARs	2006 - 2012	Electronic	MS Access, before 2009; Excel, 2009 and after	No	No		Annually	Individual ODCs around the State
			2006 - 2012	(Paper / Other)						
		3.7.1 ODC Summary Reports	2006 - 2009							

PPA DataSet Inventory - Primary Information										
Item #	Data Systems Name / Manager	Data Set Name	Years Available	Medium / Responsible Individual	Location & Type of the Data				Data Source Details	
					Database	FlatFile	Encrypted / Secure?		Update Frequency	Source(s)
4	Behavioral Risk Factor Surveillance System Data	BRFSS[8]	1984 - 2012	Electronic	MS Access	Yes, Standardized CDC format for state use.	No		Daily	Interviewers filling in CDC Questionnaire based data
5	Traumatic Brain Injury Data	TBI[9]	1996 - 2009	Electronic	Excel	No	No		Quarterly	All TN Hospitals
6	Hospital Discharge Data System (HDDS[12])	6.1 Final Inpatient	1997 - 2012	Electronic	No	Yes; in National Standard Formats - UB04 & UB92	No		None for final data sets	Hospitals, THA[23]
		6.1.1 HIPAA[10] Inpatient (De-identified Final)	1997 - 2011		No	Yes; in National Standard Formats - UB04 & UB92	No		None for final data sets	Hospitals, THA
		6.1.2 Q1-Q2 Final Inpatient Provisional	2013		No	Yes; in National Standard Format - UB04	No		Quarterly	Hospitals, THA
		6.2 Final Outpatient	1997 - 2012		No	Yes; in National Standard Formats - UB04 & UB92	No		None for final data sets	Hospitals, THA
		6.2.1 HIPAA Outpatient (De-identified Final)	1997 - 2011		No	Yes; in National Standard Formats - UB04 & UB92	No		None for final data sets	Hospitals, THA
		6.2.2 Q1-Q4 Final Outpatient Provisional	2013		No	Yes; in National Standard Format - UB04	No		Quarterly	Hospitals, THA
7	Ambulatory Surgical Treatment Center (ASTC) Data System	7.1 ASTC Final	2007 - 2012	Electronic	No	Yes; in National Standard Format - UB04 & CMS-1500	No		None for final data sets	One of three other vendors chosen by the ASTC
		7.1.1 ASTC HIPAA (De identified Final)	2007 - 2011		No	Yes; in National Standard Format - UB04 & CMS-1500	No		None for final data sets	One of three vendors chosen by the ASTC
8	Outpatient Diagnostic Center Claims Data	8.1 ODC Claims Q3	2013	Electronic	No	Yes; in National Standard Format - UB04 & CMS-1500	No		None for final data sets	One of three vendors chosen by the ODC
9	Population Estimates and Projections (PEPs)	9.1 1970-2000	1970 - 2000	(Paper / Other)	No	Yes, Standardized format for state use.	No		Annually	VS Vital Records, Demographer Contracts, and Census Data
		9.2 1970-1980	1970 - 1980	Electronic	No	Yes, Standardized format for state use.	No		Annually	VS Vital Records, Demographer Contracts, and Census Data
		9.3 1980-1990	1980 - 1990							
		9.4 1990-2020 DDPOP[19] Series	1990 - 2020							
		9.5 2000-2020 Feb '08 Revised Series	2000 - 2020							
		9.6 TN2011 - 2011 Updated Population Estimates	2011							
		9.7 TN2012 - 2012 Updated Population Estimates	2012							
		9.8 TN20102020 2010 - 2020 Updated Population Projections	2010 - 2020							
10	Births Defects Registry	Birth Defects	1999 - 2010	Electronic	No	SAS[18] Data File	No		Annually	Derived from Births (VS), HDDS, Fetal Death, & Geath

10	Births Defects Registry	Birth Defects	1999 - 2010	Electronic	No	SAS[18] Data File	No		Annually	Derived from Births (VS), HDDS, Fetal Death, & Geath
11	Early Hearing Detection & Intervention Information System (EHDI IS)	EHDI IS[20]	2007 - 2013	Electronic	MS Access	SAS Data File	No		Weekly, Monthly and Quaterly	Derived from Birth, Neometrics DB, & PTBMIS[16]
12	Pregnancy Risk Assessment Monitoring System (PRAMS)	PRAMS[21]	2007 - 2011	Electronic	CDC Based System (?) and state level MS Access	No	No		Daily	Survey answers from participating mothers; VS/VR Vital Records data sets
				Electronic	No	SAS Data File with weighted data from CDC	No		Quarterly?	Weighted survey data from CDC; weighting done on the earlier submissions from participating mothers
13	Injury Surveillance System (ISS)	ISS[11]	2001 - 2012	Electronic	No	SAS Data File	No		Quarterly and Annually	Derived from Deaths (VS), EMS[17] & HDDS

[1] VR = Vital Records
[2] VS = Vital Statistics
[3] ITP
[4] JARs = Joint Annual Reports
[5] ASTC = Ambulatory Surgical Treatment Center
[6] ACLF = Assisted Care Living Facility
[7] ODC = Outpatient Diagnostic Center
[8] BRFSS = Behavioral Risk Factors Surveillance System
[9] TBI = Traumatic Brain Injury
[10] HIPAA = Health Insurance Portability and Accountability Act
[11] ISS = Injury Surveillance System
[12] HDDS = Hospital Discharge Data System
[13] DMV = Department of Motor Vehicles
[14] AIRS = Automated Index Retrieval System
[15] KCC = Knoxville Computer Corporation
[16] PTBMIS = Patient Tracking and Billing Management Information System
[17] EMS = Emergency Medical Services
[18] SAS = Statistical Analysis System
[19] DDPOP = Not an acronym, just a Series identifier.
[20] EHDI IS = Early Hearing Detection and Intervention Information System
[21] PRAMS = Pregnancy Risk Assessment Monitoring System
[22] RW/RO = Read & Write / Read Only
[23] THA = Tennessee Hospitals Association

1.3 Summary of the Project Workflow

The workflow for registering a data request is detailed in the workflow diagrams Steps 1 – 3.

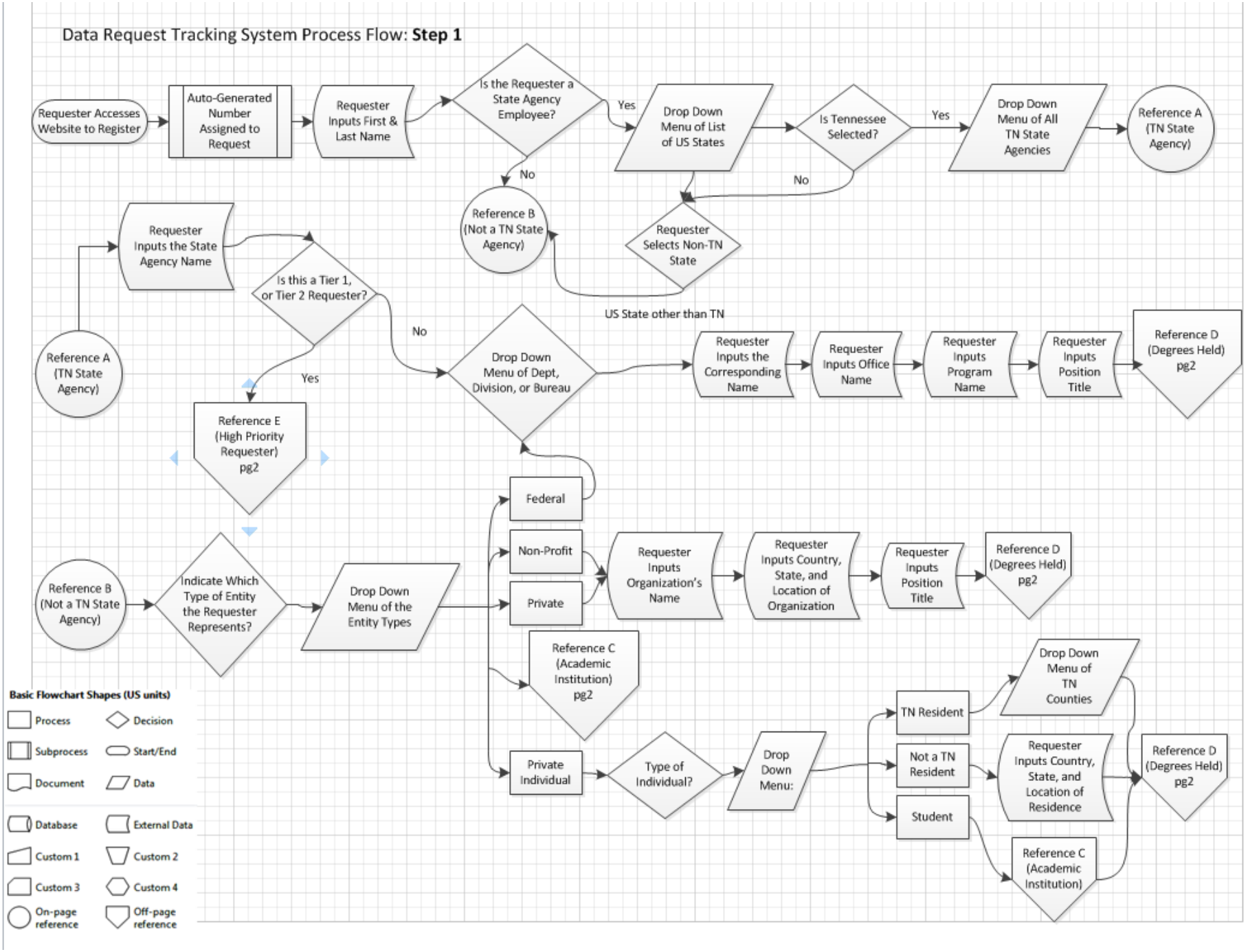
Step 1 involves identifying both the requester of the data, the nature of the request, and satisfying the required documentation based on the type of request. Step 1 involves user registration; user authentication; classifying the user according to the Tiers and therein, the level of importance and time-sensitive nature of fulfilling the data request; validating that the required documentation is provided before the request is completely processed.

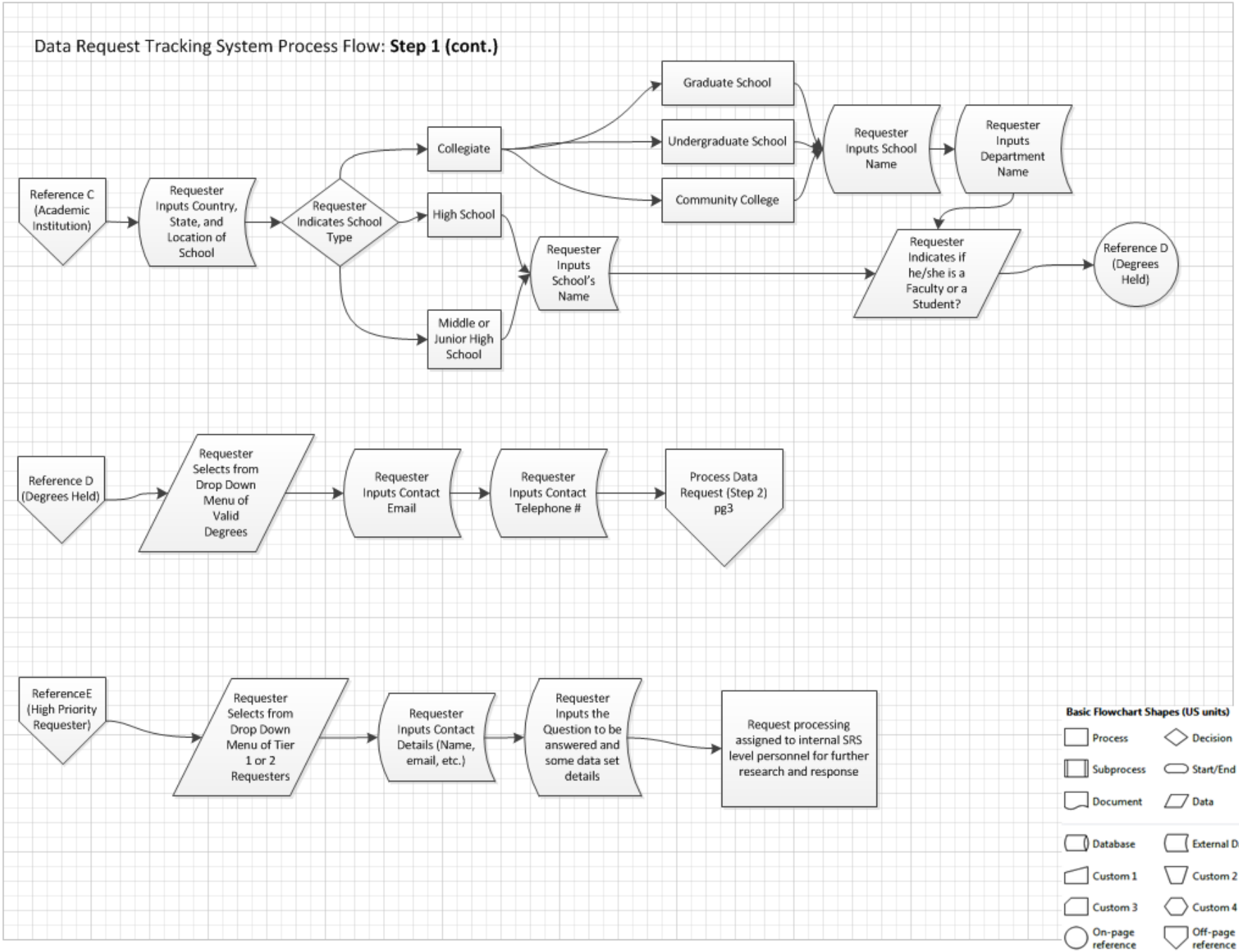
Step 2 involves classifying the desired output (type of data/output that the requester receives) involved in the data request. Classification of the data request occurs according to whether it requires aggregate vs. individual level data made available to the requester; identifiable vs. de-identified data made available to the requester; requires Institutional Review Board (IRB) approval vs. No IRB approval required; whether a confidentiality agreement or data protection plan must be included.

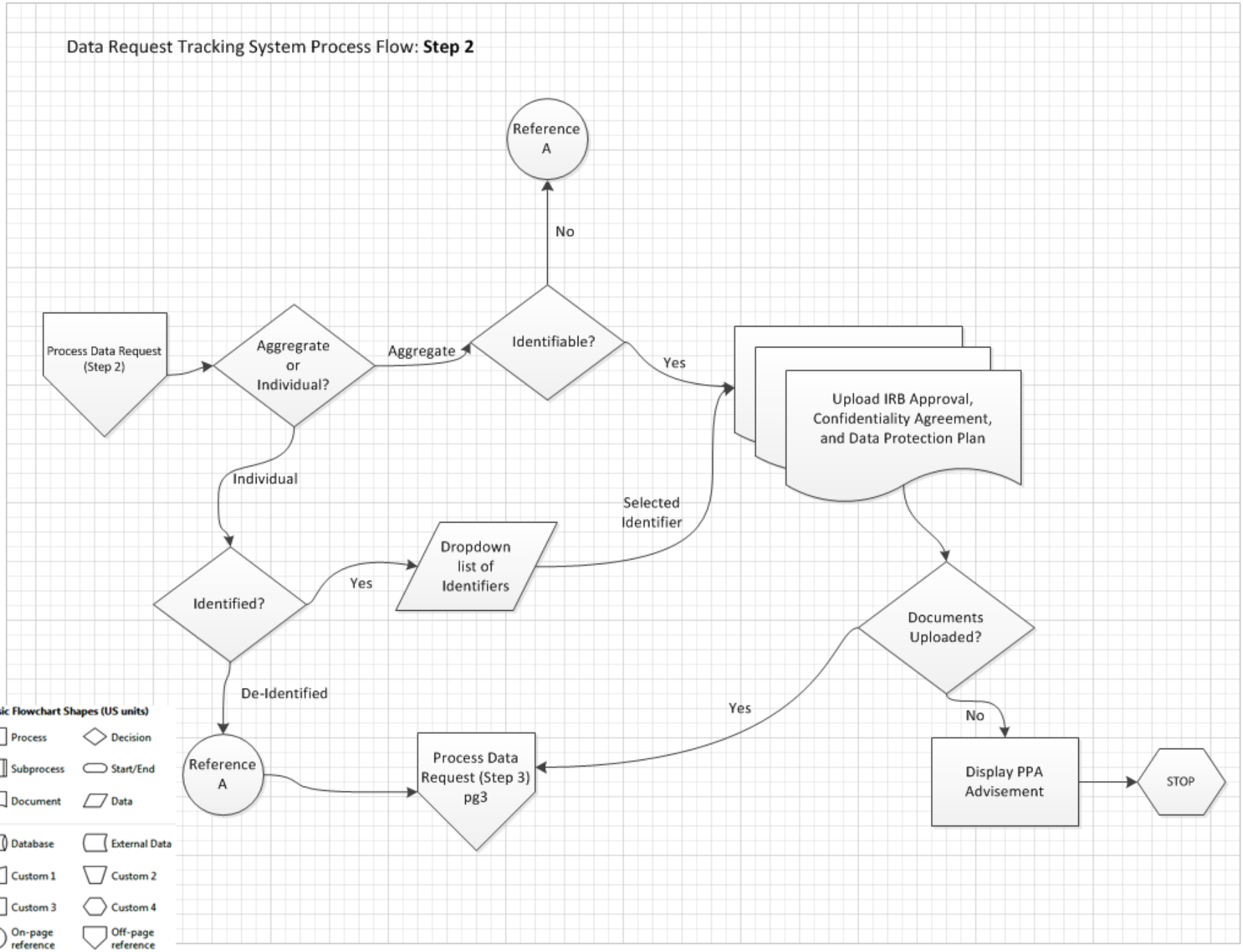
Step 3 involves further classification of the desired output from the data request according to the data available, the status of the data, and the intended consumption of the output (internal to PPA, internal to TDH, or external use). Classification of the desired output of the data request includes selecting one or more (or none) of the datasets, whether the desired output from the data request should be provisional vs. final; whether those working with the data will be internal to PPA, internal to TDH, or external to the organization and provides PPA advisement regarding policies and/or statutory requirements regarding the use and publication of the data; and whether the desired output requires no effort (raw data transfer), requires detailed analysis, requires limited analysis, or summarization.

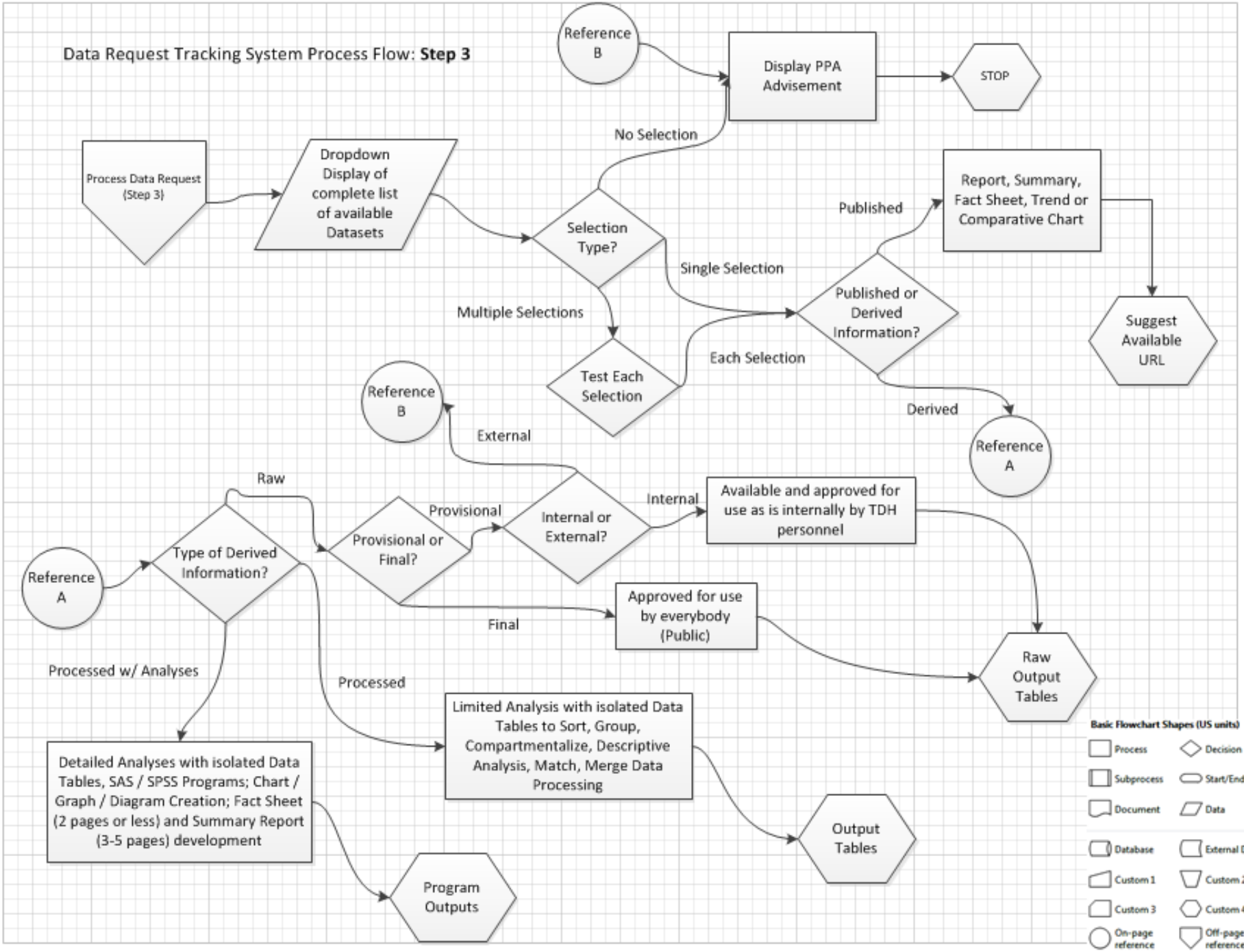
1.4 Detailed Project Workflow

Workflow diagrams are provided on the pages to follow that describe (1) the information currently captured on the paper-based form that needs to be translated to the web form; and (2) the step-by-step decision path/logic behind navigating through the web form in order to register a complete data request package using the web application.


Data Request Tracking System Process Flow: Step 1
Requester Accesses Website to Register
Auto-Generated Number Assigned to Request
Requester Inputs First & Last Name
Is the Requester a State Agency Employee?
Yes
No
Drop Down Menu of List of US States
Is Tennessee Selected?
Yes
No
Drop Down Menu of All TN State Agencies
Reference A (TN State Agency)
Reference B (Not a TN State Agency)
Requester Selects Non-TN State
US State other than TN
Requester Inputs the State Agency Name
Is this a Tier 1, or Tier 2 Requester?
Yes
No
Reference A (TN State Agency)
Reference E (High Priority Requester) pg2
Drop Down Menu of Dept, Division, or Bureau
Requester Inputs the Corresponding Name
Requester Inputs Office Name
Requester Inputs Program Name
Requester Inputs Position Title
Reference D (Degrees Held) pg2
Reference B (Not a TN State Agency)
Indicate Which Type of Entity the Requester Represents?
Drop Down Menu of the Entity Types
Federal
Non-Profit
Private
Reference C (Academic Institution) pg2
Private Individual
Requester Inputs Organization's Name
Requester Inputs Country, State, and Location of Organization
Requester Inputs Position Title
Reference D (Degrees Held) pg2
Type of Individual?
Drop Down Menu:
TN Resident
Not a TN Resident
Student
Drop Down Menu of TN Counties
Requester Inputs Country, State, and Location of Residence
Reference C (Academic Institution)
Reference D (Degrees Held) pg2
Basic Flowchart Shapes (US units)
Process
Decision
Subprocess
Start/End
Document
Data
Database
External Data
Custom 1
Custom 2
Custom 3
Custom 4
On-page reference
Off-page reference


Data Request Tracking System Process Flow: **Step 1 (cont.)**
Reference C (Academic Institution)
Requester Inputs Country, State, and Location of School
Requester Indicates School Type
Collegiate
High School
Middle or Junior High School
Graduate School
Undergraduate School
Community College
Requester Inputs School's Name
Requester Inputs School Name
Requester Inputs Department Name
Requester Indicates if he/she is a Faculty or a Student?
Reference D (Degrees Held)
Reference D (Degrees Held)
Requester Selects from Drop Down Menu of Valid Degrees
Requester Inputs Contact Email
Requester Inputs Contact Telephone #
Process Data Request (Step 2) pg3
ReferenceE (High Priority Requester)
Requester Selects from Drop Down Menu of Tier 1 or 2 Requesters
Requester Inputs Contact Details (Name, email, etc.)
Requester Inputs the Question to be answered and some data set details
Request processing assigned to internal SRS level personnel for further research and response
Basic Flowchart Shapes (US units)
Process
Decision
Subprocess
Start/End
Document
Data
Database
External Data
Custom 1
Custom 2
Custom 3
Custom 4
On-page reference
Off-page reference


Data Request Tracking System Process Flow: **Step 2**
Reference A
No
Process Data Request (Step 2)
Aggregrate or Individual?
Aggregate
Identifiable?
Yes
Upload IRB Approval, Confidentiality Agreement, and Data Protection Plan
Individual
Selected Identifier
Dropdown list of Identifiers
Identified?
Yes
Documents Uploaded?
De-Identified
Yes
No
Reference A
Process Data Request (Step 3) pg3
Display PPA Advisement
STOP
Basic Flowchart Shapes (US units)
Process
Decision
Subprocess
Start/End
Document
Data
Database
External Data
Custom 1
Custom 2
Custom 3
Custom 4
On-page reference
Off-page reference


Data Request Tracking System Process Flow: Step 3
Reference B
Display PPA Advisement
STOP
No Selection
Process Data Request (Step 3)
Dropdown Display of complete list of available Datasets
Selection Type?
Published
Report, Summary, Fact Sheet, Trend or Comparative Chart
Single Selection
Published or Derived Information?
Suggest Available URL
Multiple Selections
Test Each Selection
Each Selection
Reference B
External
Derived
Reference A
Raw
Internal
Available and approved for use as is internally by TDH personnel
Provisional
Internal or External?
Provisional or Final?
Reference A
Type of Derived Information?
Final
Approved for use by everybody (Public)
Raw Output Tables
Processed w/ Analyses
Processed
Limited Analysis with isolated Data Tables to Sort, Group, Compartmentalize, Descriptive Analysis, Match, Merge Data Processing
Detailed Analyses with isolated Data Tables, SAS / SPSS Programs; Chart / Graph / Diagram Creation; Fact Sheet (2 pages or less) and Summary Report (3-5 pages) development
Output Tables
Program Outputs
Basic Flowchart Shapes (US units)
Process
Decision
Subprocess
Start/End
Document
Data
Database
External Data
Custom 1
Custom 2
Custom 3
Custom 4
On-page reference
Off-page reference

1.5 Description of User and Functional Interfaces

The data request tool that we envision would consist of distinct elements as follows:

1. **Registration of a Data Requester and creation of a secure login for the Requester**. There shall be no limitations to the number of data requests and user registrations.
2. **Logging-in to the Data Request system via the World Wide Web (WWW) and should be compatible with all the available web browsers** from Microsoft (Internet Explorer), Apple (Safari), Google (Chrome), and Mozilla (Firefox).
3. **Creation of a new Data Request package** along with requisite supporting documents.
4. **Data Request administration and permission groups for employees responding to data requests**:
    a. Request Overview and classification of the Requester as being one of the four following categories:
        i. **Public end-users** with ability to register for an account and submit a data request.
        ii. **Frontline employees** of the State of Tennessee with ability to view and respond to data request assigned to the employee or the team that the employee has been assigned to work in. This category of users generally produces aggregation of raw data and outputs in the form of tables, graphs/charts, summaries and fact sheets.
        iii. **Mid-level employees** of the State of Tennessee with ability to view, respond, and assign tasks for fulfilling the data request to teams under him/her. This category of users may handle patient level record requests and might analyze datasets to match, merge, tabulate, etc.; produce health statistics and health summary reports (>5 pages).
        iv. **High-level employees** with ability to view, respond, edit work of others, audit and validate results of data request, assign tasks to fulfill data requests to teams and individuals, communicate with other teams for additional data set dependencies, and review complex research data requests that may involve patient-level information as well as other complex analytical functions that may involve multiple datasets.
    b. **Tracking of distribution of Data Request including**, internal validation, and data request fulfillment task assignments and schedules.
    c. Data Request Sign-Off and Submission
    d. **Communication of Data Request as needed** to different categories of data request handlers (as described above).
    e. Deletion of Data Request
    f. Data Request History and user transaction log.

The creation of a new Data Request solicits requester information in terms of:

- Affiliation of the requester
- Requester title, name, degrees held, etc.
- Data Request sponsor, if any
- Data Request keywords
- Assignment of an internal Request ID.
- Assembly of the Data Request package that includes all the required supporting documentation depending on the nature of the data request and the outputs desired by the requester. An example follows:
    - Attach one or more supporting documents (e.g., IRB approval letter, Conflict of Interest statement, HIPAA Consent, etc.) already prepared by the user and is available on the requester's computer.
    - Instructions for filling out a specific form required for data request
    - Information about the data requester, and the details of level of access to the state owned data (e.g., analytic files, aggregate confidential data, etc.)
    - Information about additional personnel who would have access to the data request output.

- Main contact information for the Data Request
- Information about the Data Request funding sources - federal, state, etc.
- Information about any third party contractor involvement
- Information about waivers and any special populations that might be involved
- Information about Datasets and Limited Datasets
- Affirmations (for such things as protecting any data supplied by TDH to the user for research purposes only) and signatures.

The tool must allow a detailed web conversation with the data requester to help him/her assemble a data request package that includes all necessary and context-dependent supporting information, advise and help the user in providing all the required information in order to submit a valid request. In case of any errors (e.g., incorrect information submitted, lack of proper authorization, etc.) during the process, the "Data Request Administration" functionality identified above should provide all the necessary navigational steps and documentation required to assist the user complete the data request package and get it registered along with a valid ID in the PPA at TDH. There shall be no specific time limit in which the user is expected to complete the creation of a valid data request package. The user will be assisted by a two-way conversation between him/her and the PPA personnel via email and/or messaging to furnish the required supporting documents, signatures or authorizations, etc. Thus the process of creating a data request package might be interrupted, user inputs saved, and finished at a later time. The system thus manages and maintains data request package creation histories, package ownership and completion status, and ages any partially completed package as required.

The process of validating a data request package and subsequent acceptance, registration for processing, and data fulfillment is exemplified by a use case illustrated in the next section.

2. Product Features

2.1 Product Perspective

PPA's fulfillment of data requests requires that we consider the interests of our customers (both internal and external) and the quality of our services as far as data request handling is considered. The idea is to provide a web-based application to structure IT-related activities and streamline the interactions of data requesters with personnel within PPA.

Currently, PPA uses an in-house developed system based on MS Access to manually update and track incoming data requests. This system has been found to be grossly inadequate for tracking requests that are mostly coming via telephone and email mostly on an adhoc basis. Furthermore, the burden of tracking and improving the system involved with fulfilling the data requests has been delegated to personnel with little or no formal training in workflow analyses that are charged with doing the backend analyses work and/or collecting and refining the raw data as needed. Finally, the end user (esp., the public sending requests via a web interface on a TDH web site) often gets engaged in multiple and prolonged conversations with the PPA personnel in order to satisfy the additional documentation required for the fulfillment of a data request. Even after all these manual steps are undertaken by the PPA personnel, documentation regarding the fulfilled data request is fragmented and inconsistent. The inconsistencies in the process lead to wavering fulfillment and/or customer satisfaction levels with the quality of the service provided by PPA.

Yet another shortcoming of our current system is in terms of tracking the different categories of users from individuals and academic researchers to personnel from the Governor's, Legislator's, and Commissioner's offices or from public/private institutions like legal and health-related organizations.

The process flow diagram in section 1.4 describes the complete process of providing a new data request for the PPA. The complete process entails three distinct steps, each involving a number of sub-steps as detailed by the process flow diagram. Step 1 consists of allowing all intended users of the system an online access whereby the user is prompted to answer a number of closed-ended questions to determine the user's affiliations, intent and use for requested data, and to classify the user into one of the four tiers as has been determined by the members of the new data reform committee at the PPA. At the end of Step 1, the data requestor will be classified into one of the four tiers as follows:

- Tier 1 - Personnel from Commissioner's Office, Governor's Office, Media, or Legislator/Congressman's Office
- Tier 2 - Personnel for Public Release of Information and Office of General Counsel
- Tier 3 - Internal or External Agency Employee
- Tier 4 - Academic and/or Qualified research personnel and General Public Citizen

In addition to classification of the data requester, Step 1 also consolidates all the required and any additional documentation required for providing the needed data request. The additional documentation might come in many forms - anywhere from details of data requestor's affiliation (internal agency person, external researcher, etc.) to specification of specific type of needed information (e.g., individual, raw, processed, etc.). Any time during the classification process of the data requester if it deems that the data requester will be denied to request information, a proper PPA advisement is given to the data requestor in order that the requestor may understand and prepare for the kind of additional documentation and/or certification required for the request of the needed data.

When once all the necessary sub-steps of Step 1 are satisfied the requestor is cleared for Step 2 and Step 3 described below.

Step 2 consists of figuring out the required additional documentation (if any) for processing the data request. This information is usually in the form of IRB (Institutional Review Board) related documentation and protected Personal Health Information (PHI) related Confidentiality Agreements. Once again, if the requester cannot provide the needed additional information the data request process is aborted with a PPA advisement to the requester.

Finally, when all the Step 1 and Step 2 sub-processes are done, the user is directed to Step 3 where specific PPA data sets in question, the type of data processing required , and the required PPA personnel whose services are required for data request processing are determined. A final decision as to how the request needs to be processed will then be available and suitable workflow for the data request processing is initiated.

2.2 Product Features

The primary functionality required for the Data Request Tool may be summarized as follows and their detailed descriptions are given in section 3. The expected product features are:

1. The tool front-end is primarily the World Wide Web and should be compatible with all the available web browsers from Microsoft (Internet Explorer), Apple (Safari), Google (Chrome), and Mozilla (Firefox).
2. The tool is enabled for use via a valid log-in via pre-registered UserID and a Password (with expiration options and configurable security levels).
3. The primary graphical user interface (GUI) should provide only needed and closed-end questionnaire in order to determine both the data requester's credentials and his/her data request details. User prompts must be based on suitable skip logic and only valid user answers must be solicited and no room given for open-ended and narrative statements.
4. All required user documentation shall be provided via attachments with pre-determined file types (doc, docx, pdf, etc.) and will be collected and stored separately for each requester and managed for privacy and confidentiality in a secure manner.
5. Time stamps of all required and significant user inputs will be collected as the requester enters data and approximate request fulfillment times be posted as required and logged for user convenience and/or reporting of incomplete documentation and other such timely episodes.
6. All requests for data will be archived and all user data collected will have pre-determined expiry dates and permanent archival mechanisms.
7. The tool must be equipped with suitable mechanisms for approvals, tiered process level tags, process completion statuses, and activity logs for feedback bot to the data requester and the PPA data management team.
8. The tool must be enabled with an ability for internal PPA personnel to act as proxies to a specified user (as the circumstances dictate) and should be able to consolidate multiple users to represent a single entity such as an academic institution or a government body such as a Commissioner's office. In this sense, multiple users from the same entity may represent one consolidated user (when required) and may exercise the user privileges of a single user at a known tier.
9. For those cases where a data request may have to go through multiple internal approvals, a mechanism must be provided by the tool software to communicate with the principals for approvals and manage the requisite signatures with time stamps (electronic) collection in order to orderly process the data request.
10. A mechanism must be made available to capture user activities and inputs at the front-end and transfer it as an email or text message to the administrator for use in team meetings for approvals, additional documentation corrections, etc.
11. There will be accommodations for an administrative level user for the Data Request

tool and shall be given special privileges to set up user accounts, privileges, and other administrator functions.

12. At the time of this writing, the number of user licenses required at the level of a public citizen is “unlimited” and at the administrator and/or internal PPA levels, the required number of concurrent user licenses is estimated to be 2 or 3 with an assumption that each concurrent user license may support up to 3-5 users. More details may be worked out at a later date with respect to the number of concurrent user licenses.

2.3 User Classes and Characteristics

As mentioned earlier there are users that are customers (public and users/data requesters in Tiers 1, 2, 3, and 4); and there are users that are employees fulfilling requests (frontline, mid-level, and high-level). Tiers 1 and 2 users are expected to be personnel from the highest levels of state government such as the Governor’s Office, Commissioner’s Office, the Legislature, or the Media. Tier 3 users are expected to be from the legal entities or the public agencies (or individuals) involved with Information Request under the state or federal law. Tier 4 users comprise of all others including research personnel from the academia and other public/private institutions/organizations, non-profits, etc. No specific educational level or expertise is expected of the users and the user might be highly educated and trained or a novice user from a middle or high school working on some school project. Hence the Step 1 of the data request process is designed for an elaborate discovery of the type of data requester including the characteristics of the data he/she is requesting.

The following tables identify the time-sensitive deadlines that are deemed reasonable for fulfilling a data request:

Type	Requester Type	Reasonable Response Timeframe (Standard, Not Urgent Request)
Tier 1	Commissioner’s Office	1 – 3 Business Days
Tier 1	Governor’s Office	1 – 3 Business Days
Tier 1	Media Request	1 – 3 Business Days
Tier 1	Legislator/Congressperson	1 – 3 Business Days
Tier 2	Public Release of Information Request	2 – 7 Business Days
Tier 2	Office of General Counsel (OGC)	2 – 7 Business Days

Type	Requester Type	Reasonable Response Timeframe (Standard, Not Urgent Request)
Tier 3	Agency Employee (Internal)	5 – 8 Business Days
Tier 3	Agency Employee (External)	5 – 8 Business Days

Type	Requester Type	Reasonable Response Timeframe (Standard, Not Urgent Request)
Tier 4	Academic and/or Qualified Research	8 – 14 Business Days
Tier 4	General Public Citizen Request	8 – 14 Business Days

To the extent possible, the front-end user interface must be designed to lead the data requester in every possible way to request information that is meant to be accessible for the type of user he/she is determined to be by the software while also limiting any attempt to input data which is either irrelevant and/or extraneous for the specific answers to be provided for the user. To recap, there will be three levels of users within TDH who are also employees. They are:

a. **Frontline employees** of the State of Tennessee with ability to view and respond to data request assigned to the employee or the team that the employee has been assigned to work in. This category of users generally produces aggregation of raw data and outputs in the form of tables, graphs/charts, summaries and fact sheets.
b. **Mid-level employees** of the State of Tennessee with ability to view, respond, and

assign tasks for fulfilling the data request to teams under him/her. This category of users may handle patient level record requests and might analyze datasets to match, merge, tabulate, etc.; produce health statistics and health summary reports (>5 pages).

c. **High-level employees** with ability to view, respond, edit work of others, audit and validate results of data request, assign tasks to fulfill data requests to teams and individuals, communicate with other teams for additional data set dependencies, and review complex research data requests that may involve patient-level information as well as other complex analytical functions that may involve multiple datasets.

2.4 Example of a Data Request Use Case

The example of a Principal Investigator (PI) as a data requester is used in the following example illustrating the creation of a typical Data Request package:

- PI accesses the online web application for Data Request and registers himself/herself first with a personal User ID and a Password along with a few security questions and answers in order to help retrieval and/or resetting of a forgotten UserID and Password.
- The PI logs in for creating a new Data Request package.
- Upon successful login, the PI is presented with a number of PDF documents that describes various datasets available for request and a number of help (or training) documents that describe the creation process, user responsibilities with the data obtained, and any required authorizations for the use of data. These documents may be downloaded and printed (or saved) on the PI's computer for later use.
- Next, the PI is asked to input the PI's personal and project information along with the Access level (raw, aggregate, analytical, etc.) required for the use of data.
- If the PI has associated who might be looking at the data requested, their personal and project information must also be provided in separate input forms.
- Next, the funding information for the project is solicited and depending on the nature of funding (e.g., Federal, State, Private Trust, etc.) additional information will be solicited. There could be multiple sources for funding and each source needs to be identified as required.
- More information sources are needed to be identified if the PI uses human subjects, has biological samples, and/or uses any hospital based data. These details are acquired by individual and separate screens (or Appendices) as needed.
- If the PI has any involvement with a third party contractor and/or personnel working for the State (or TDH), additional information about those people will be required to be shown in the Data Request package.
- Next, specific questions regarding the datasets requested will need to be answered in terms of checking off the respective checkboxes associated. In addition, any other required ancillary datasets from other state agencies must also be identified at a level of detail necessary.
- Finally, the PI is asked to affirm the Data Request package and obtain any necessary signatures and other supporting documents as required before saving and closing the newly created Data Request package.

Further processing of the newly created Data Request package for validating and classifying the user, the requested dataset and the type of functions required to operate on the dataset, and the PPA personnel that are needed to work on fulfilling the request, etc. are elaborated briefly in section 3 below.

2.5 Operating Environment

The Data Request tool software is expected to operate in a typical home or office PC, laptop, or a tablet computer environment equipped with standard Internet browsers like IE, Safari, Chrome,

or Firefox. Mobile application compatibility is preferred but not required at this time.

On the server side which supports the Internet browser-based client, it is envisioned that some sort of light database and/or web server would be needed with the usual security services, data management software environment for the user and system files, administrative services to enable user log-ins, set up account privileges, etc. No special application interfaces are expected for the operation of the Data Request tool software environment. 100% server-side hosting by the vendor on a cloud is expected and preferred for the Data Request tool. Any required customization of the cloud-based product to fit the needs of PPA and TDH will be handled by the personnel from PPA and ITSD as required.

The product customization and maintenance would be directed by the IT services provided by the Tennessee state government at TDH and would involve working with the Office of Information Resources (OIR) and the Office of the Information Technology Services (OITS). The software vendor would be expected to work with OIR and OITS personnel as required during the product installation (*if necessary*) and scheduled maintenance periods.

2.6 Design and Implementation Constraints

At this writing no special constraints have been identified for the Data Request tool software that might limit the options available for the development of software. The language to be used by the users is expected to be English; however, a Spanish language front-end interface is preferred but not required for Spanish speaking users seeking health information from TDH.

As far as any security and software maintenance issues are concerned, it is expected that TDH would be responsible for providing those services as needed by the ITSD at TDH after the delivery of the complete software package to TDH.

2.7 User Documentation

Both user level and administration level documentation must be provided in order to train, operate, and maintain the software as required.

2.8 Assumptions and Dependencies

No special assumptions and dependencies are expected for the regular installation, operation, and maintenance of the Data Request tool software.

3. System Functional Requirements

This section identifies the system features in terms of detailed functional requirements enabling the three step process identified earlier for accepting valid data requests from the four tiers of users for this product.

3.1 User Qualification and Classification

3.1.1 Description and Priority

Qualifying and classifying a data requester is an extremely important and is a high priority function for the Data Request tool. The specific category that the user would belong to determines both the kind of data he/she may request and the response time within which the request would be fulfilled. Any Tier 1 or 2 user request must be so identified appropriately and must be tracked within the system so that the response time of 1-7 business days would be met by the responsible person in PPA and the data request fulfillment process becomes visible to all requisite PPA personnel.

Data requests coming from Tier 3 and 4 users must also be processed in a timely manner (5-14 business days) but more importantly, the data request must be accompanied by suitable additional documentation as required depending on the specific data request. To this end, the process steps 2 and 3 would be of higher priority than mere classification of the data requester and all the additional user inputs as required for the data request must be satisfied. In addition, the longer time required to process approvals, etc. must also be tracked and a complete log of the data request and fulfillment activity be documented for a possible audit later on.

3.1.2 Stimulus/Response Sequences

Figure B-1 illustrates the detailed process flow of Step 1 activities. The main stimulus/response requirements are described in the table below in terms of a user dialog:

Item #	Stimulus	Response
1	User log-in	Known user - Start Step 1; Unknown user - Prompt for Registration into the system
2	User identification details	Classify according to TN State Agency employee or not; get more data requester details in either case
3	TN State Agency employee details	Accept details and go to Step 2
4	Non-TN State Agency person	Ask for further details about Organization type, Residency etc.
5	Non-TN State Agency person details	Classify the user as Individual (student, etc.) or some federal or other US state agency/organization employee; accept requester details and go to Step 2

3.1.3 Functional Requirements

Any errors encountered during the above user dialog must be clarified to the user with simple and straightforward error messages and/or in-actions that signify to the user that he/she has made a mistake in the data entry. As far as possible, the data requester details are input using closed-end single or multiple check pull-down menus so that entry errors are minimized. Any free form text input must be minimized and should be restricted to personal educational qualifications, agency or personal location address, and such.

Any user input requiring more than a couple of lines must be solicited using attachments with options for the document type (doc, docx, or pdf).

As far as possible, lists of all internal TN State named agencies, organizations that may use health data, etc. will be provided to the vendor to populate drop-down menus in order that the use of “other” category and free format text input is made to be minimal.

3.2 Data Request Classification

3.2.1 Description and Priority

The classification of the incoming data request is as important (if not more) as classifying the data requester and is also of a high priority. Incorrect data request data classification may incur loss of time in fulfilling the request and/or not being fully able to satisfy the data requester with the right amount and type of health data being provided. Without correct classification, the required additional data request attachments and approvals might be missed thus causing delays and improper handling of the data request processing.

The incorrect processing and delays especially apply to IRB-related requests and any identifiable data embedded in the data requests. The process sub-steps required here must be carefully planned and executed with full support and cooperation of PPA personnel in order that all required documents, data sets, approvals, etc. are carefully monitored, documented and verified before making the software fully deployable.

3.2.2 Stimulus/Response Sequences

Section 1.4 illustrates the detailed process flow of Step 2 activities. The main stimulus/response requirements are described in the table below in terms of a user dialog:

Item #	Stimulus	Response
1	Data Request type is Aggregate	Check if it is identifiable and if so start asking for supporting documents such as IRB approval, Confidentiality Agreement, etc.; If not identifiable go to Step 3
2	Data Request type is Individual	Provide drop-down list of known identifiers for selection; Upon specific selection start asking for supporting documents such as IRB approval, Confidentiality Agreement, etc.
3	Additional documents uploaded	Accept details and go to Step 3
4	Missing additional documentation	Display PPA Advisement and stop further processing

3.2.3 Functional Requirements

Just like in 3.1.3 any errors encountered during the above user dialog must be clarified to the user with simple and straightforward error messages and/or in-actions that signify to the user that he/she has made a mistake in the data entry. As far as possible, the data requester details are input using closed-end single or multiple check pull-down menus so that entry errors are minimized. No free form text input is envisioned for this step, but rather a specific list of documents for upload.

As far as possible, lists of all identifiers, names of required IRB documents, etc. will be provided to the vendor to populate drop-down menus in order that the use of “other” category and free format text input is made to be minimal.

3.3 Data Request Processing

3.3.1 Description and Priority

The processing of the validated data request is the most important part and tracking the process is the reason behind the design of a structured process to deal with TDH data requests. While the details of specific functions involved in this step is detailed elsewhere (section 3.3.3 below), this is the highest priority step that may involve many levels of PPA personnel depending on the specifics of the data request.

A high level description of the sub-processes involved in this step is to identify the data set(s) involved, decide on the analyses that may be required, identify manipulations on raw data, etc. and further determine the PPA personnel that need to be assigned to the tasks involved before generating the required answers for the incoming data query. It is also important to identify any readily obtainable answers from a previously answered request so that regeneration and repetition of work is avoided while still fulfilling the data request. Thus this step involves decision-making, logical analyses, and gathering and/or generating of reports and facts as needed.

3.3.2 Stimulus/Response Sequences

Section 1.4 illustrates the detailed process flow of Step 3 activities. The main stimulus/response requirements are described in the table below in terms of PPA internal stimuli and responses:

Item #	Stimulus	Response
1	Set of all PPA Datasets available for selection	Check if it is a single selection or multiple selections; Test each selection for a published or derived information
2	Published Dataset Information	Provide links to published report, fact sheet, table, etc. on the TDH web sites as required
3	Derived data from selected Dataset(s)	Determine further dataset type, analyses required, provisional or final dataset for use, etc.
4	Provisional dataset selected by an external data requester	Display PPA Advisement and stop further processing
5	Approved Dataset for use by the requester	Check if further analyses on the data are required; If so, process the data further; Generate required reports and outputs for dissemination to the requester.

3.3.3 Functional Requirements

The specific functions involved in this step are listed below:

- Decide upon the required PPA Dataset(s) for processing a valid data request
- Classifying the selected Dataset in terms of whether a previously generated report, fact sheet, etc. is available for immediate dissemination
- Deciding whether the dataset needed has provisional or final version of the data and then checking the user credentials for being internal or external for further processing consideration.
- Figuring out the different types of dataset operations required as the need may be - selecting specific subsets of data from a dataset, analyzing specific rows and columns of a dataset table to generate new tables, charts, maps, etc.
- Generating actual reports, etc. and disseminating the processed results back to the data requester.

- Housekeeping operations and logging of all data processing activities and storing generated results in known results data repositories for later use.

Any processing errors are to be handled by PPA personnel and are not visible to the data requester. Any further advisement for the data requester due to any processing errors must be handled separately on a case-by-case basis.

4. External Interface Requirements

4.1 User Interfaces

No special GUI standards are identified or this Data Request tool other than the standard web interfaces found in the four popular web browsers mentioned earlier, i.e., Internet Explorer, Safari, Chrome, and Firefox.

4.2 Hardware Interfaces

No special hardware devices and interfaces are identified for this Data Request tool other than the standard keyboard, mouse and other pointing devices used in surfing the Internet (e.g., touch tablets, arrow keys, etc.).

4.3 Software Interfaces

The Data Request tool is expected to be hosted on PC clients on the client-side and some Enterprise OS-based server in the server environment provided by the vendor on a cloud with any required customization by OIR and ITSD personnel. No other special OSes are expected on the client side other than Windows, Apple iOS, and any other OS that supports the web browsers mentioned earlier (IE, Safari, Chrome, and Firefox).

4.4 Communications Interfaces

No special communications interfaces are identified are identified with the use of the Data Request tool. Standard HTTP and FTP provided on both server and client side resources would be more than enough to support the operations of this Data Request tool.

5. Other Nonfunctional Requirements

5.1 Performance Requirements

Since the Data Request system's front-end user interface is based on WWW, the solicitation of all required input data from the user must happen at normal conversational speeds on the web, irrespective of the hardware and the browser the user may be using. Any and all excessive processing times at the back-end must be conveyed (intimated) to the user suitably to avoid user frustrations and any detailed request for additional information must be via previously generated documents in standard formats such as MS Word, Adobe PDF, JPG/PNG/etc.-formatted graphics and image files, etc.

5.2 Safety Requirements

No special safety requirements have been determined for the installation and running of the Data Request tool.

5.3 Security Requirements

Forgotten usernames, passwords and security features must all be handled by proper administrative functions and user feedbacks within the web application. No secure dataset names and attributes are made visible to the WWW user and the system must guard against unintended display of any personal data.

5.4 Software Quality Attributes

No special software quality attributes have been determined for the installation and running of the Data Request tool.

Appendix A: Glossary

GUI	Graphical user Interface
HIPAA	Health Insurance Portability and Accountability Act
IE	Internet Explorer
IRB	Institutional Review Board
IT	Information Technology
ITSD	Information Technology Systems Division
JPG	Joint Photographers expert Group
MS	Microsoft
OIR	Office of Information Resources
OITS	Office of the Information Technology Services
OPM	Office of Performance Management
OS	Operating System
PDF	Portable Document Format
PHI	Personal Health Information
PNG	Portable Network Graphics
PPA	Policy, Planning, and Assessment
TDH	Tennessee Department of Health
WWW	World Wide Web